

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/821,687

DATE: 07/27/2001

TIME: 10:46:08

Input Set : A:\81356162.app

Output Set: N:\CRF3\07272001\I821687.raw

3 <110> APPLICANT: MIKOSHIBA, KATSUHIKO  
 4 MIZUTANI, AKIHIRO  
 6 <120> TITLE OF INVENTION: RNA-BINDING PROTEIN  
 8 <130> FILE REFERENCE: 081356/0162  
 10 <140> CURRENT APPLICATION NUMBER: 09/821,687  
 C--> 11 <141> CURRENT FILING DATE: 2001-07-03  
 13 <150> PRIOR APPLICATION NUMBER: JP 2000-299812  
 14 <151> PRIOR FILING DATE: 2000-09-29  
 16 <160> NUMBER OF SEQ ID NOS: 11  
 18 <170> SOFTWARE: PatentIn Ver. 2.1  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 483  
 22 <212> TYPE: DNA  
 23 <213> ORGANISM: Mus musculus  
 25 <220> FEATURE:  
 26 <221> NAME/KEY: CDS  
 27 <222> LOCATION: (1)..(483)  
 29 <400> SEQUENCE: 1  
 30 att gaa att gtt ttt gct aag cca cca gat cag aag agg aaa gaa aga 48  
 31 ile Glu ile Val Phe Ala Lys Pro Pro Asp Gln Lys Arg Lys Glu Arg  
 32 1 5 10 15  
 34 aaa gct cag agg caa gca gca aag aat caa atg tat gat gat tac tac 96  
 35 Lys Ala Gln Arg Gln Ala Ala Lys Asn Gln Met Tyr Asp Asp Tyr Tyr  
 36 20 25 30  
 38 tat tat ggt cca cct cat atg cct ccc cca aca aga ggt cga ggg cgt 144  
 39 Tyr Tyr Gly Pro Pro His Met Pro Pro Pro Thr Arg Gly Arg Gly Arg  
 40 35 40 45  
 42 gga ggt aga ggt ggc tat gga tat cct cca gat tat tat gga tac gaa 192  
 43 Gly Gly Arg Gly Gly Tyr Gly Tyr Pro Pro Asp Tyr Tyr Gly Tyr Glu  
 44 50 55 60  
 46 gat tat tat gat tat tat ggt tat gat tac cat aac tat cgt ggt gga 240  
 47 Asp Tyr Tyr Asp Tyr Tyr Gly Tyr Asp Tyr His Asn Tyr Arg Gly Gly  
 48 65 70 75 80  
 50 tat gaa gat cca tac tat ggt tat gaa gat ttt caa gtt gga gct aga 288  
 51 Tyr Glu Asp Pro Tyr Tyr Gly Tyr Glu Asp Phe Gln Val Gly Ala Arg  
 52 85 90 95  
 54 gga agg ggt ggt aga gga gca agg ggt gct gct cca tcc aga ggt cgt 336  
 55 Gly Arg Gly Gly Arg Gly Ala Arg Gly Ala Ala Pro Ser Arg Gly Arg  
 56 100 105 110  
 58 ggg gct gct cct ccc cgt ggt aga gcc ggt tat tca cag aga gga ggc 384  
 59 Gly Ala Ala Pro Pro Arg Gly Arg Ala Gly Tyr Ser Gln Arg Gly Gly  
 60 115 120 125  
 62 cct gga tca gca aga ggc gtt cgc ggt gcg aga gga ggt gcc caa caa 432  
 63 Pro Gly Ser Ala Arg Gly Val Arg Gly Ala Arg Gly Gly Ala Gln Gln  
 64 130 135 140  
 66 caa aga ggc cgc ggg gga aaa ggg gtc gag gcc ggt cct gac ctg tta 480  
 67 Gln Arg Gly Arg Gly Gly Lys Gly Val Glu Ala Gly Pro Asp Leu Leu

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/821,687

DATE: 07/27/2001

TIME: 10:46:08

Input Set : A:\81356162.app

Output Set: N:\CRF3\07272001\I821687.raw

```

68 145          150          155          160
70 caa
71 Gln
74 <210> SEQ ID NO: 2
75 <211> LENGTH: 161
76 <212> TYPE: PRT
77 <213> ORGANISM: Mus musculus
79 <400> SEQUENCE: 2
80 Ile Glu Ile Val Phe Ala Lys Pro Pro Asp Gln Lys Arg Lys Glu Arg
81 1 5 10 15
83 Lys Ala Gln Arg Gln Ala Ala Lys Asn Gln Met Tyr Asp Asp Tyr Tyr
84 20 25 30
86 Tyr Tyr Gly Pro Pro His Met Pro Pro Thr Arg Gly Arg Gly Arg
87 35 40 45
89 Gly Gly Arg Gly Gly Tyr Gly Tyr Pro Pro Asp Tyr Tyr Gly Tyr Glu
90 50 55 60
92 Asp Tyr Tyr Asp Tyr Tyr Gly Tyr Asp Tyr His Asn Tyr Arg Gly Gly
93 65 70 75 80
95 Tyr Glu Asp Pro Tyr Tyr Gly Tyr Glu Asp Phe Gln Val Gly Ala Arg
96 85 90 95
98 Gly Arg Gly Gly Arg Gly Ala Arg Gly Ala Ala Pro Ser Arg Gly Arg
99 100 105 110
101 Gly Ala Ala Pro Pro Arg Gly Arg Ala Gly Tyr Ser Gln Arg Gly Gly
102 115 120 125
104 Pro Gly Ser Ala Arg Gly Val Arg Gly Ala Arg Gly Gly Ala Gln Gln
105 130 135 140
107 Gln Arg Gly Arg Gly Gly Lys Gly Val Glu Ala Gly Pro Asp Leu Leu
108 145 150 155 160
110 Gln
113 <210> SEQ ID NO: 3
114 <211> LENGTH: 3452
115 <212> TYPE: DNA
116 <213> ORGANISM: Mus musculus
118 <220> FEATURE:
119 <221> NAME/KEY: CDS
120 <222> LOCATION: (154)..(1839)
122 <400> SEQUENCE: 3
123 gcggttttgc tgccagcggc gtgagcttcg gccgccattt tacaacagct ccactcgcgc 60
125 cggacacagg gagcagcgag cagcggttcc ccgcccagccg acccggtcgg acggattcct 120
127 cggccccagc cccgcgggga gatctctgga aac atg gct aca gaa cat gtt aat 174
128 Met Ala Thr Glu His Val Asn
129 1 5
131 gga aat ggt act gaa gag ccc atg gat act act tca gca gtt atc cat 222
132 Gly Asn Gly Thr Glu Glu Pro Met Asp Thr Thr Ser Ala Val Ile His
133 10 15 20
135 tca gaa aat ttt cag aca ttg ctt gat gct ggt tta cca cag aaa gtt 270
136 Ser Glu Asn Phe Gln Thr Leu Leu Asp Ala Gly Leu Pro Gln Lys Val
137 25 30 35
139 gct gaa aaa cta gat gaa att tac gtt gca ggg cta gtt gca cat agt 318

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/821,687

DATE: 07/27/2001

TIME: 10:46:08

Input Set : A:\81356162.app

Output Set: N:\CRF3\07272001\I821687.raw

140	Ala	Glu	Lys	Leu	Asp	Glu	Ile	Tyr	Val	Ala	Gly	Leu	Val	Ala	His	Ser	
141	40					45					50					55	
143	gat	tta	gat	gaa	aga	gct	atc	gaa	gct	tta	aaa	gag	ttc	aat	gaa	gac	366
144	Asp	Leu	Asp	Glu	Arg	Ala	Ile	Glu	Ala	Leu	Lys	Glu	Phe	Asn	Glu	Asp	
145				60						65					70		
147	ggc	gca	ttg	gca	gtg	ctt	caa	cag	ttt	aaa	gac	agt	gat	ctc	tct	cat	414
148	Gly	Ala	Leu	Ala	Val	Leu	Gln	Gln	Phe	Lys	Asp	Ser	Asp	Leu	Ser	His	
149				75					80					85			
151	gtt	cag	aac	aaa	agt	gcc	ttt	tta	tgt	gga	gtc	atg	aag	act	tac	agg	462
152	Val	Gln	Asn	Lys	Ser	Ala	Phe	Leu	Cys	Gly	Val	Met	Lys	Thr	Tyr	Arg	
153			90					95					100				
155	cag	aga	gaa	aaa	cag	ggg	acc	aaa	gta	gca	gac	tct	agt	aaa	gga	cca	510
156	Gln	Arg	Glu	Lys	Gln	Gly	Thr	Lys	Val	Ala	Asp	Ser	Ser	Lys	Gly	Pro	
157		105					110					115					
159	gat	gag	gca	aag	att	aag	gca	ctt	ttg	gaa	aga	aca	ggc	tac	aca	ctt	558
160	Asp	Glu	Ala	Lys	Ile	Lys	Ala	Leu	Leu	Glu	Arg	Thr	Gly	Tyr	Thr	Leu	
161	120					125					130					135	
163	gat	gtg	act	aca	ggt	cag	agg	aag	tat	gga	gga	cca	cct	cca	gat	tcc	606
164	Asp	Val	Thr	Thr	Gly	Gln	Arg	Lys	Tyr	Gly	Gly	Pro	Pro	Pro	Asp	Ser	
165				140						145					150		
167	gtt	tat	tca	ggt	cag	cag	cct	tct	gtt	ggc	act	gag	ata	ttt	gtg	ggg	654
168	Val	Tyr	Ser	Gly	Gln	Gln	Pro	Ser	Val	Gly	Thr	Glu	Ile	Phe	Val	Gly	
169				155						160				165			
171	aag	atc	ccc	aga	gat	ctg	ttt	gag	gat	gag	ctt	gtt	cca	tta	ttt	gag	702
172	Lys	Ile	Pro	Arg	Asp	Leu	Phe	Glu	Asp	Glu	Leu	Val	Pro	Leu	Phe	Glu	
173			170					175					180				
175	aaa	gct	gga	cct	ata	tgg	gat	ctt	cgt	tta	atg	atg	gat	ccg	ctc	act	750
176	Lys	Ala	Gly	Pro	Ile	Trp	Asp	Leu	Arg	Leu	Met	Met	Asp	Pro	Leu	Thr	
177		185				190						195					
179	ggt	ctc	aac	aga	ggt	tat	gcg	ttt	gtc	act	ttt	tgt	aca	aaa	gaa	gca	798
180	Gly	Leu	Asn	Arg	Gly	Tyr	Ala	Phe	Val	Thr	Phe	Cys	Thr	Lys	Glu	Ala	
181	200					205					210				215		
183	gca	caa	gag	gct	gtt	aaa	ctg	tat	aat	aat	cat	gaa	att	cgt	tcc	ggg	846
184	Ala	Gln	Glu	Ala	Val	Lys	Leu	Tyr	Asn	Asn	His	Glu	Ile	Arg	Ser	Gly	
185				220						225					230		
187	aag	cac	att	ggt	gtc	tgc	atc	tca	gtt	gcc	aac	aat	agg	ctt	ttt	gtg	894
188	Lys	His	Ile	Gly	Val	Cys	Ile	Ser	Val	Ala	Asn	Asn	Arg	Leu	Phe	Val	
189				235					240					245			
191	ggc	tcg	att	cct	aag	agt	aaa	acc	aag	gag	cag	att	ctt	gag	gaa	ttt	942
192	Gly	Ser	Ile	Pro	Lys	Ser	Lys	Thr	Lys	Glu	Gln	Ile	Leu	Glu	Glu	Phe	
193			250					255					260				
195	agc	aaa	gtg	aca	gag	ggt	ctc	aca	gat	gtc	att	tta	tac	cac	caa	cct	990
196	Ser	Lys	Val	Thr	Glu	Gly	Leu	Thr	Asp	Val	Ile	Leu	Tyr	His	Gln	Pro	
197		265				270					275						
199	gat	gac	aag	aaa	aaa	aac	aga	ggc	ttt	tgc	ttt	ctt	gaa	tat	gaa	gat	1038
200	Asp	Asp	Lys	Lys	Lys	Asn	Arg	Gly	Phe	Cys	Phe	Leu	Glu	Tyr	Glu	Asp	
201	280					285					290				295		
203	cac	aaa	aca	gct	gcc	cag	gca	aga	cgt	agg	cta	atg	agt	ggt	aaa	gtc	1086
204	His	Lys	Thr	Ala	Ala	Gln	Ala	Arg	Arg	Arg	Leu	Met	Ser	Gly	Lys	Val	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/821,687

DATE: 07/27/2001

TIME: 10:46:08

Input Set : A:\81356162.app

Output Set: N:\CRF3\07272001\I821687.raw

```

205          300          305          310
207 aaa gtc tgg gga aat gtt gga act gtt gag tgg gct gat cct att gaa 1134
208 Lys Val Trp Gly Asn Val Gly Thr Val Glu Trp Ala Asp Pro Ile Glu
209          315          320          325
211 gat cct gat cct gaa gtt atg gca aag gta aaa gtg ctg ttt gta cgc 1182
212 Asp Pro Asp Pro Glu Val Met Ala Lys Val Lys Val Leu Phe Val Arg
213          330          335          340
215 aac ctt gcc aac acg gta aca gaa gaa att tta gaa aag tca ttt agt 1230
216 Asn Leu Ala Asn Thr Val Thr Glu Glu Ile Leu Glu Lys Ser Phe Ser
217          345          350          355
219 cag ttt ggg aaa ctg gaa cga gta aag aag cta aaa gat tat gct ttc 1278
220 Gln Phe Gly Lys Leu Glu Arg Val Lys Lys Leu Lys Asp Tyr Ala Phe
221 360          365          370          375
223 att cat ttt gat gag aga gat ggt gct gtc aag gct atg gaa gaa atg 1326
224 Ile His Phe Asp Glu Arg Asp Gly Ala Val Lys Ala Met Glu Glu Met
225          380          385          390
227 aat ggt aaa gac ttg gag gga gaa aat att gaa att gtt ttt gct aag 1374
228 Asn Gly Lys Asp Leu Glu Gly Glu Asn Ile Glu Ile Val Phe Ala Lys
229          395          400          405
231 cca cca gat cag aag agg aaa gaa aga aaa gct cag agg caa gca gca 1422
232 Pro Pro Asp Gln Lys Arg Lys Glu Arg Lys Ala Gln Arg Gln Ala Ala
233          410          415          420
235 aag aat caa atg tat gat gat tac tac tat tat ggt cca cct cat atg 1470
236 Lys Asn Gln Met Tyr Asp Asp Tyr Tyr Tyr Tyr Gly Pro Pro His Met
237          425          430          435
239 cct ccc cca aca aga ggt cga ggg cgt gga ggt aga ggt ggc tat gga 1518
240 Pro Pro Pro Thr Arg Gly Arg Gly Arg Gly Gly Arg Gly Gly Tyr Gly
241 440          445          450          455
243 tat cct cca gat tat tat gga tac gaa gat tat tat gat tat tat ggt 1566
244 Tyr Pro Pro Asp Tyr Tyr Gly Tyr Glu Asp Tyr Tyr Asp Tyr Tyr Gly
245          460          465          470
247 tat gat tac cat aac tat cgt ggt gga tat gaa gat cca tac tat ggt 1614
248 Tyr Asp Tyr His Asn Tyr Arg Gly Gly Tyr Glu Asp Pro Tyr Tyr Gly
249          475          480          485
251 tat gaa gat ttt caa gtt gga gct aga gga agg ggt ggt aga gga gca 1662
252 Tyr Glu Asp Phe Gln Val Gly Ala Arg Gly Arg Gly Gly Arg Gly Ala
253          490          495          500
255 agg ggt gct gct cca tcc aga ggt cgt ggg gct gct cct ccc cgt ggt 1710
256 Arg Gly Ala Ala Pro Ser Arg Gly Arg Gly Ala Ala Pro Pro Arg Gly
257          505          510          515
259 aga gcc ggt tat tca cag aga gga ggc cct gga tca gca aga ggc gtt 1758
260 Arg Ala Gly Tyr Ser Gln Arg Gly Gly Pro Gly Ser Ala Arg Gly Val
261 520          525          530          535
263 cgc ggt gcg aga gga ggt gcc caa caa caa aga ggc cgc ggg gga aaa 1806
264 Arg Gly Ala Arg Gly Gly Ala Gln Gln Gln Arg Gly Arg Gly Gly Lys
265          540          545          550
267 ggg gtc gag gcc ggt cct gac ctg tta caa tga agactgactt gctattgtgg 1859
268 Gly Val Glu Ala Gly Pro Asp Leu Leu Gln
269          555          560

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/821,687

DATE: 07/27/2001

TIME: 10:46:08

Input Set : A:\81356162.app

Output Set: N:\CRF3\07272001\I821687.raw

```

271 gattacacca gaagcttgca gtggagtaat ggtaaggaaa atcaagcaac cttaaataatc 1919
273 tcggatgtat aggagcatat tctattgcag aagaccctcc tatgaagatc atggaatcaa 1979
275 atacgggaca ttgaactaat acttggaactt tgttatgaat ttctttaaca attttctctg 2039
277 cagtgcgaagt tattaaacta aagctactct attttccaaa tgtgttccaa aaaaatcctt 2099
279 cataacttct agcatggtat ctttaataaag aataaagttg ttctctttaa aaaatctgct 2159
281 ctaagtagat ttttcccctc ttttttaatt aaggatctca gcagtgggtat tctgaaatat 2219
283 tctcttgaat ttgtgcattt aaattttatt gcagtgatac agatgccact gttggtaccc 2279
285 ttaaattttt atttctgctc accaagggtta atcatgattg tctatatctt ttttatagta 2339
287 atcacttttg aattgtgttc agatgcagtt tcaggtgtaa tcatcagagc tggtagtca 2399
289 ggcattccag atagtgttc ttttcagaac ctttttaag ggttggttaa ctacctcagt 2459
291 agcagaggat tgaactatac cctgtctgta ctgtacatag aaaatctttg tagataaaag 2519
293 caaggttgt taaatatgat atgagggtaa gattttaata taccaaatgt aacattctta 2579
295 gttgccttta gtttcagagg cttgtaagac ttctcatga ccatcataac aggccttgct 2639
297 tttgtcgtat tttgtggctg aaaaagcagc cttgcttctt cagatattgt agttatttgg 2699
299 atgtataata gtttagcaag atgttacttt tgtaagacat cagatgttca aaaaaaaagt 2759
301 gcatccgaac ttgtactaaa tactgcagtg tccctttata aaaagtcaga ctaaaactga 2819
303 caattgtaca gcaaagcctg acatttggat attttgaagt tttttcataa atcatagaaa 2879
305 ttagtatatg gctgtagttt agcttttttag gtaaaaggta tgtttcatta gtgcatttgt 2939
307 tattgtgat cactataaaa atgtgaatca gctttccatt tcttatgcag gtcatgataa 2999
309 cttgtagaat agagtacaat catttgtgct atgtttttaa ttttctaaag caccttgatg 3059
311 acaatgagtg ttcagtggtg aagcctctc tattgaatca cctcaaaaaa atttttttgc 3119
313 caagtccata gttgatagct taaagtcaaa agtaaaatta tagtttaagt aggacttggg 3179
315 gtaaagaaac accctcccc tccccaaaag ggatactgca gttctatcac acaccagta 3239
317 ggcaccacga tgacgatcag agctagacct gactaagggt ttatacacac cagttcccca 3299
319 gtaaagtcaa atttaacaag aaaattagac atgtcatatg ttcaaaatgc tcatggcaaa 3359
321 caatcatttt gcattcctgc aaataaaatt gttttatacc agcttggact taaccaggct 3419
323 gaacttgcta atcggatecc cgggtaccga gct 3452

```

326 &lt;210&gt; SEQ ID NO: 4

327 &lt;211&gt; LENGTH: 561

328 &lt;212&gt; TYPE: PRT

329 &lt;213&gt; ORGANISM: Mus musculus

331 &lt;400&gt; SEQUENCE: 4

```

332 Met Ala Thr Glu His Val Asn Gly Asn Gly Thr Glu Glu Pro Met Asp
333   1             5             10             15
335 Thr Thr Ser Ala Val Ile His Ser Glu Asn Phe Gln Thr Leu Leu Asp
336           20           25           30
338 Ala Gly Leu Pro Gln Lys Val Ala Glu Lys Leu Asp Glu Ile Tyr Val
339         35         40         45
341 Ala Gly Leu Val Ala His Ser Asp Leu Asp Glu Arg Ala Ile Glu Ala
342       50       55       60
344 Leu Lys Glu Phe Asn Glu Asp Gly Ala Leu Ala Val Leu Gln Gln Phe
345   65       70       75       80
347 Lys Asp Ser Asp Leu Ser His Val Gln Asn Lys Ser Ala Phe Leu Cys
348       85       90       95
350 Gly Val Met Lys Thr Tyr Arg Gln Arg Glu Lys Gln Gly Thr Lys Val
351     100     105     110
353 Ala Asp Ser Ser Lys Gly Pro Asp Glu Ala Lys Ile Lys Ala Leu Leu
354     115     120     125
356 Glu Arg Thr Gly Tyr Thr Leu Asp Val Thr Thr Gly Gln Arg Lys Tyr

```

Use of n and/or Xaa has been detected in the Sequence Listing.  
 Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/821,687

DATE: 07/27/2001

TIME: 10:46:09

Input Set : A:\81356162.app

Output Set: N:\CRF3\07272001\I821687.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:501 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8

L:518 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9